

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on:

December 8, 2001, 18:56:41 ; Search time 1735.73 Seconds
(without alignments)
24528.455 Million cell updates/sec

Title:

US-08-153-397A-1

Perfect score:

3962

Sequence:

1 CGGGCTGAGACTGGGTGA.....AAAAAACCAGGATTC 3962

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gdext 1.0

Searched:

11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estfun:*

2: em_esthum:*

3: em_estin:*

4: em_estom:*

5: em_estpl:*

6: em_estba:*

7: em_estro:*

8: em_estov:*

9: em_hic:*

10: gb_estl:*

11: gb_est2:*

12: gb_hic:*

13: gb_gss:*

14: em_gss_fun:*

15: em_gss_hum:*

16: em_gss_inv:*

17: em_gss_pln:*

18: em_gss_pro:*

19: em_gss_rod:*

20: em_gss_vrt:*

21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1720	43.4	2633	12	BC006836	Mus muscu
2	960	24.2	1010	10	AL539517	AL539517
3	815.2	20.6	1019	10	AL528663	AL528663
4	814.6	20.6	987	10	AL528664	AL528664
5	715.8	18.1	770	11	BI257820	BI257820
6	684	17.3	799	11	BG912321	BG912321
7	580.4	17.2	715	10	BE744806	BE744806
8	672.4	17.0	719	11	BF345815	BF345815
9	672	17.0	744	11	BI193181	BI193181
10	662.6	16.7	718	11	BG696171	BG696171
11	660.2	16.7	772	10	AL043251	AL043251
12	655.6	16.5	912	10	BE304984	BE304984

13	654.6	16.5	652	11	BF944646
14	653.8	16.5	877	10	BE249888
15	652	16.5	957	10	BE737252
16	650.8	16.4	683	11	BE752521
17	650.4	16.4	685	11	BG403848
18	646.4	16.3	787	11	BF338761
19	646	16.3	1144	11	BF527983
20	643.6	16.2	898	11	BE898332
21	640.8	16.2	801	10	BE735635
22	639.6	16.1	1018	11	BE899403
23	636.2	16.1	759	11	BF344637
24	635	16.0	921	11	BF792714
25	632.2	16.0	1009	10	BE613181
26	631.2	15.9	759	11	BG913935
27	629.2	15.9	711	11	BG541910
28	628.4	15.9	1150	11	BF344537
29	627.4	15.8	652	10	AL047539
30	627.2	15.8	804	11	BG696424
31	626.6	15.8	977	11	BG386352
32	624	15.7	732	11	BF339561
33	621	15.7	831	10	BE513178
34	620.8	15.7	626	10	AW376875
35	619.6	15.6	1037	11	BF529368
36	613	15.5	680	11	BG818241
37	603.6	15.2	649	10	BE304585
38	602.4	15.2	1325	11	BF966734
39	594.8	15.0	645	10	BG395129
40	591.2	14.9	904	11	BG289311
41	588.4	14.9	614	10	BE305175
42	586.8	14.8	600	10	AW629699
43	583.8	14.7	611	11	BG820294
44	583	14.7	596	11	BG831101
45	580.8	14.7	922	11	BG116520

ALIGNMENTS

RESULT 1

BC006836

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC006836 2633 bp mRNA
1, clone IMAGE:3598142, mRNA.
BC006836
BC006836.1 GI:14711754
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.
Direct Submission
Submitted (27-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalona@bcm.tmc.edu
Villalona, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAP Plate: 16 Row: 0 Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis

FEATURES

```

1. .2633
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="3598142"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
483 a 805 c 825 g 520 t
BASE COUNT
ORIGIN

```

	Query Match	43.48;	Score 1720;	DB 12;	Length 2059;
	Best Local Similarity	79.98;	Pred. No. 9.8e+309;		
	Matches 2200;	Conservative	0; Mismatches 310;	Indels 244;	Gaps 5;
QY	277	ACAGATGCTGCCCCACCACCCCCTTAGCGCCGAGGGATCAGSAGACTAGSGACACGACGAGCCCC	336		
DB	121	AGAGATGCTGTCCTTCCATAGCACAGAGGATCCAGACGATGGGACAGGACCC	180		
QY	337	TGTCACTTTTACTGCTGTG---CTCTTGTCGAAGTGGAGATGCTGACATGAAGGAC	393		
DB	181	TCTCATCTCTACTGCTGCTACTCTTGTCGACAAATTGGAGATGCTGACATGAAGGAC	240		
QY	394	ATTTTGATCTGCACAGTGCCTATGTCCTGGCCATGCAGACCGGACATCCACAGA	453		
DB	241	ATTTTGACCTGCCAAGTGCCTATGTCCTGGCCATGCAGACCGGACATCCACAGA	300		
QY	454	GTGACATCTCTGTCTTCAGATCTCTGTGTCAGATTCACCTGCCGCCGCCACAGACAGCTTG	513		
DB	301	CGCATATCTGTCTCCAGCTCTCTGTGTCGACTCTACCGCTGCCGCCGCCACAGACGCTGG	360		
QY	514	AGACAGCTGACGGGGATGGGCCCTGGTGGCCCGCAGGTCGGTGTTCCTCCAAGAGGAGG	573		
DB	361	AAGCAGTGATGGAGATGGGCTTGGTGGCCCTGCAGGGCCCTGTGTTCCCAAAGAAGAGG	420		
QY	574	AGTACTTCAGTGGATCAACAGACTCCACACTTCGTCGTGGTCTCGTGGGACCCAGGGAC	633		
DB	421	AGTACTTCAGTGGATCAACAGACTCCACACTTCGTCGTGGTCTCGTGGGACCCAGGGAC	480		
QY	634	GGCATCCGGGGGCTGGGCAAGGAGTTCCTCCGAGACTACCGGCTGGCTACTCCCGG	693		
DB	481	GGCATCTGGGGGCTGGGCAAGGATTCCTCCGAAGCTATCGGTTGGCTACTCCCCAG	540		
QY	694	ATGGTCGCGCTGGATGGGCTGGAAGGACCGCTGGGCTCAGGAGTGATCTCAGGCAATG	753		
DB	541	ATGGCCCGCTGGATGGGCTGGAAGGACCGCTGGGACAGGAGGTATTCGGGTAAAG	600		
QY	754	AGGACCTGAGGAGTGGTCTGAAGSACCTTGGGCCCCCATGGTTGCCACATGCTTC	813		
DB	601	AGGATCCCGGGAGTAGTCTGAAGGACCTTGGACCCCATGGTTGGCCCGGCTGGTCC	660		
QY	814	GCTTCTACCCCGGCTGACCGGCTCATGAGTCTCTGTGCGGGTAGAGCTCATGGCT	873		
DB	661	GCTTCTACCCCGGCTGACCGGATCATGAGTCTCTGTCTCCTGGGTGGAGCTCATGGCT	720		
QY	874	GCCTCTGAGGAGGATGCTCTGTCTTACACCGCCCTGTGGGACAGACAATGTAATTAT	933		
DB	721	GCTCTGGCG-----	730		
QY	934	CTGAGGCGGTGTACTCAACAGCTCCACCTATGACGACATACCGTGGCGGAGCTGCAGT	993		
DB	731	-----GGGCTGCAGT	740		
DB					

```

Db 1698 CCTATGCTGTCGCTGCACTGCCGCCAGGGGGGGTGGGGATGGGCCGCCAGAGTGGATT 1757
QY 2134 TCCTCGATCTCGACTCCGTTCAAGGAGAGCTTGGCGAGGGCCAGTTTGGGAGGTGC 2193
Db 1758 TCCTCGGTCACGGCTCCGCTTCAAGGAGAGCTTGGCGAGGGCCATTTGGGAGGTAC 1817
QY 2194 ACCTGTGTAGGTGCGACAGCCCTCAAGATCTGTGTCAGTCTGTGATTTCCCTTAAATGTC 2253
Db 1818 ACCTGTGTAGGTGCGACAGCCCTCAAGATCTGTGTCAGTCTGTGATTTCCCTTAAATGTC 1877
QY 2254 GTAAGGACACCTTTCTGCTAGTCTCAAGATCTTACGCCAGATGCCACCAAGATG 2313
Db 1878 ACAAGGACACCTTTCTGCTAGTCTCAAGATCTTCCCGCCAGATGCCACCAAAATG 1937
QY 2314 CCAGCTTCTCTCTTCCAGGAATGATTTCTCTGAAGAGAGTGAACATCATGTGCGAGGC 2373
Db 1938 -----CCAGGAATGATTTCTCTGAAGAGAGTGAACATCATGTGCGAGGC 1979
QY 2374 TCAAGACCCCAACATCATTTCCGCTGCTGGGCTGTGTGTCAGGACGACCCCTCTGCA 2433
Db 1980 TGAAGGACCAACATCATCTCCGCTCTGGTGTGTGTGTCAGGATGACCCCTCTGCA 2039
QY 2434 TGATTACTGACTACATGGAGAGCGGACCTCAACAGTTCCTCAGTGGCCACAGCTGG 2493
Db 2040 TGATCAGACTACATGGAATGGGATCTGAACAGTTCCTCAGTGGCCGCGAGCTGG 2099
QY 2494 AGGACAGGACGCGGAGGGGCCCTTGGGACGGGAGGCTGCGAGGGGCCACCATCA 2553
Db 2100 AGACAGGCGCACTAGGGGCTCTTGGGACACAGATCTGACGAGGGGCCACCAATCA 2159
QY 2554 GTACCAATGTCTGTGATGTGGACCCAGATCGCTCCCGCATCGCTATCTGGCCA 2613
Db 2160 GCTACCTATGTGTACAGCTGGGGCCAGATCGCTTGCATGCGTTATCTGCCA 2219
QY 2614 CACTCAACTTTGATATCGGAGCTGGCCACGGGAACCTCTATCTGGGACTATTACCGTG 2733
Db 2220 CACTCAAAATCGCGACTTGGCATGAGTCGGAATCTACGCTGGGGATTTATACCGTG 2339
QY 2734 TGCAGGGCGGCGAGTGTGCCATCCGCTGGATGGCTGGGAGTGCTATCTCATGGGA 2793
Db 2340 TCCAGGGCGGCGGTGTGCCATCAGTGTGATGGCTGGGAGTGCTATCTCATGGGA 2399
QY 2794 AGTTACAGACTCGAGTGTGGGCTTTGGTGTGACCTGTGGGAGTGTCTGATGC 2853
Db 2400 AGTTACACACAGCCAGTGTGGGCTTCCGAGTGACCTGTGGGAGTGTCTGATGC 2459
QY 2854 TCTGTAGGCGGCGGCTTGGGAGCTACCCAGCAGGAGTATCGAGAGCGGGGG 2913
Db 2460 TCTGAGTGTCCAGGCTTGGGAGCTTACAGATGAGCAGTATCGAGAAATCGCGGG 2519
QY 2914 AGTTCTCCGGACAGGCGGCGAGTGTACCTGTCCCGGCGCTGTGCTGCGGCGAG 2973
Db 2520 AGTTCTTCCAGGACAGGCGGCGAGTGTACTTGTCCAGGCGACCCGCTGCCACAGA 2579
QY 2974 GCCTATATGATGTGTGGTGTCTGGAGCGGAGTGTCTGAGCAGCGACAC 3027
Db 2580 CCCTGTATGATGTATGCTCGGTTTGGAGCGGAGCGGCGGAGCGGCGGCG 2633

```

RESULT 2

```

AL539517 1010 bp mRNA EST 16-FEB-2001
LOCUS AL539517 LTI_FL013_FBrn1 Homo sapiens cDNA clone CS0DF033Y015 5
DEFINITION prime, mRNA sequence.
ACCESSION AL539517
VERSION AL539517.1 GI:12868804
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 1010
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DF033Y015"
/clone_lib="LTI_FL013_FBrn1"
/dev_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"
/lab_host="DH108"
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(df) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

FEATURES
source

BASE COUNT 191 a 321 c 298 g 199 t 1 others
ORIGIN

Query Match 24.2%; Score 960; DB 10; Length 1010;
Best Local Similarity 99.5%; Pred. No. 5.8e-168;
Matches 983; Conservative 1; Mismatches 1; Indels 3; Gaps 2;

QY 2500 AGCAGCGGAGGGGGCCCTTGGGACGGGAGGCTGGCGAGGGGCCACCATCAGCTACC 2559
Db 1 AGCAGCGGAGGGGGCCCTTGGGACGGGAGGCTGGCGAGGGGCCACCATCAGCTACC :60
QY 2560 CAATGCTGCTGATGTGGCAGCCAGATCGCTTCCGGCATCGCTATCTGGCCACACTCA 2619
Db 61 CAATGCTGCTGATGTGGCAGCCAGATCGCTTCCGGCATCGCTATCTGGCCACACTCA 120
QY 2620 ACTTTGTACATCGGGACCTGGCCACCGGAACTGCCTAGTTGGGAAATTTCCACCATCA 2679
Db 121 ACTTTGTACATCGGGACCTGGCCACCGGAACTGCCTAGTTGGGAAATTTCCACCATCA 180
QY 2680 AATCCGAGACTTTGGCATGAGCCGGAACCTCTATCTGGGGACTATTACCGTGTGCAGG 2739
Db 181 AATCCGAGACTTTGGCATGAGCCGGAACCTCTATCTGGGGACTATTACCGTGTGCAGG 240
QY 2740 GCGGGCACTGCTGCCCATCCGCTGGATGGCTTGGAGTGCACTCTCATGGGGAAGTTCA 2799
Db 241 GCGGGCACTGCTGCCCATCCGCTGGATGGCTTGGAGTGCACTCTCATGGGGAAGTTCA 300
QY 2800 CGACTCGAGTACGCTGTGGGCTTTGGTGTGACCTGTGGAGGTGCTGATCTCTGTA 2859
Db 301 CGACTCGAGTACGCTGTGGGCTTTGGTGTGACCTGTGGAGGTGCTGATCTCTGTA 360
QY 2860 GGGCCACAGCCCTTTGGGCACTCACCGACGAGCTCATCGAAGACGGGGGAGTCTT 2919
Db 361 GGGCCACAGCCCTTTGGGCACTCACCGACGAGCTCATCGAAGACGGGGGAGTCTT 420
QY 2920 TCCGGACAGCGGCGGAGGTGTACCTGTCCCGGCGGCTGTGCTCCCGCGAGGCTAT 2979
Db 421 TCCGGACAGCGGCGGAGGTGTACCTGTCCCGGCGGCTGTGCTCCCGCGAGG--CTAT 478
QY 2980 ATGAGCTGATGCTTGGTGTGAGCGGAGTCTGAGCAGCGACCATCTTTTCCACAGC 3039
Db 479 ATGAGCTGATGCTTGGTGTGAGCGGAGTCTGAGCAGCGACCATCTTTTCCACAGC 538

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1010)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 1010

Db 59 TGATTTTACACTTATATATTGCTCCTTGTGAGGAAATTTTAATTATATGAWTAGG 2

RESULT 4
AL528664

LOCUS AL528664 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD001YK21 5
DEFINITION prime, mRNA sequence.
ACCESSION AL528664
VERSION AL528664.1 GI:12792157
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 987)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. 987
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DD001YK21"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 186 a 263 c 332 g 200 t 6 others
ORIGIN

Query Match 20.6%; Score 814.6; DB 10; Length 987;
Best Local Similarity 98.7%; Pred. No. 5e-141;
Matches 848; Conservative 4; Mismatches 3; Indels 4; Gaps 3;

QY 277 AGAGATGCTGCCCCACCCCTTAGGCCCGAGGGATCAGGAGCTATGGACACGAGGCC 336
|||
Db 90 AGARATGCTGCCCCACCCCTTAGGCCCGAGGGATCAGGAGCTATGGACACGAGGCC 149
|||
QY 337 TGTCATCTTTACTGCTGCTCTTGTGTGGCAAGTGAGATGCTGACATGAAGGACATT 396
|||
Db 150 TGTATCTTTTACTGCTGCTCTTGTGTGGCAAGTGAGATGCTGACATGAAGGACATT 209
|||
QY 397 TTGATCTCTGCAAGTGCCTATGCTCCTCGGCATGCAAGACCGACCATCCACAGAGTG 456
|||
Db 210 TTGATCTCTGCAAGTGCCTATGCTCCTCGGCATGCAAGACCGACCATCCACAGAGTG 269
|||
QY 457 ACATCTCTGCTTCAGCTCTTGCTCAGATTCCTACTGCGCCCGCCACAGAGGTTGGAGA 516
|||
Db 270 ACATCTCTGCTTCAGCTCTTGCTCAGATTCCTACTGCGCCCGCCACAGAGGTTGGAGA 329
|||
QY 517 GCAGTGACGGGGATGGGGCTGTCGCCCGCAGGGTGGTGTTCCTCCAAGGAGGAGGT 576
|||
Db 330 GCAGTGACGGGGATGGGGCTGTCGCCCGCAGGGTGGTGTTCCTCCAAGGAGGAGGT 387
|||
QY 577 ACTTTCAGGTGGATCTACACGACTCCACTGCTGCTCTGTTGGGACCCAGGACGCC 636
|||
Db 388 ACTTTCAGGTGGATCTACACGACTCCACTGCTGCTCTGTTGGGACCCAGGACGCC 447
|||

/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 Kb. Library prepared by Life
Technologies."
144 a 207 c 262 g 157 t

BASE COUNT
ORIGIN

Query Match 18.1%; Score 715.8; DB 11; Length 770;
Best Local Similarity 97.3%; Pred. No. 1e-122;
Matches 750; Conservative 0; Mismatches 17; Indels 4; Gaps 2;
QY 229 GAAGTGGCTATTCACTGAGCGATGGGTTGGACTTGAAGGATGCCAAGAGATGCTGCC 288
DB 1 GAAGTGGCTATTCACTGAGCGATGGGTTGGACTTGAAGGATGCCAAGAGATGCTGCC 60
QY 289 CCACACCCCTTAGGCCCGGAGGATGAGGATGATGGACAGAGGCCCTGTCTATCTTTAC 348
DB 61 CCACACCCCTTAGGCCCGGAGGATGAGGATGATGGACAGAGGCCCTGTCTATCTTTAC 120
QY 349 TGTGCTGCTCTTGTGGCAAGTGGAGATGCTGACATGAAGGACATTTGATCCTGCCA 408
DB 121 TGTGCTGCTCTTGTGGCAAGTGGAGATGCTGACATGAAGGACATTTGATCCTGCCA 180
QY 409 AGTGGCGCTATGCCCTGGGCATGCCAGGACCGGACCATCCACAGCATGACATCTCTGCTT 468
DB 181 AGTGGCGCTATGCCCTGGGCATGCCAGGACCGGACCATCCACAGCATGACATCTCTGCTT 240
QY 469 CCAGCTCTCTGGTCAGATTCCTACCTGCCGCCGCCACAGCAGGTTGGAGACATGACGGGG 528
DB 241 CCAGCTCTCTGGTCAGATTCCTACCTGCCGCCGCCACAGCAGGTTGGAGACATGACGGGG 300
QY 529 ATGGGGCTTGTGTCGCCCGCAGGCTGGTGTTCCTCCAAAGGAGGAGTACTTGCAGGTGG 588
DB 301 ATGGGGCTTGTGTCGCCCGCAGGCTGGTGTTCCTCCAAAGGAGGAGTACTTGCAGGTGG 360
QY 589 ATCTACAGCATCCACTGCTGGTCTGGTGGGACCCAGGAGCGCATGCCGGGGGCC 648
DB 361 ATCTACAGCATCCACTGCTGGTCTGGTGGGACCCAGGAGCGCATGCCGGGGGCC 420
QY 649 TGGCAAGGAGTCTCCCGGAGCTACCGGCTGCGTTACTCCCGGATGTCGCCCTGGA 708
DB 421 TGGCAAGGAGTCTCCCGGAGCTACCGGCTGCGTTACTCCCGGATGTCGCCCTGGA 480
QY 709 TGGGCTGGAAGGACCGCTGGGTCAGGAGTGTATCTCAGGCAATGAGGACCTGAGGGAG 768
DB 481 TGGGCTGGAAGGACCGCTGGGTCAGGAGTGTATCTCAGGCAATGAGGACCTGAGGGAG 540
QY 769 TGGTCTGAAGACCTTGGGCCCCCCTGTTGGCCGAGTGTGCTTCTACCCCGGG 828
DB 541 TGGTCTGAAGACCTTGGGAAACCCCTGTTGGCCGAGTGTGCTTCTACCCCGGG 600
QY 829 CTGACCGGTCTATG---AGTGTCTCTGCGGGTAGAGCTCTATGGCTGCTCTGAGGG 885
DB 601 CTGACCGGTCTATGAGCAGTCTGTACTGCGGGTAGAGCTCTATGGCTGCTCTGAGGG 660
QY 886 ATGGACTCTCTGTCTACACCCCTGTTGGGCGCAGACATGATTTATCTGAGGCCGTGT 945
DB 661 ATGGACTCTCTGTCTACACCCCTGTTGGGCGCAGACATGATTTATCTGAGGCCGTGT 719
QY 946 ACCTCAACGACTCCACCTATGACGACATACCGTGGGGGGGACTGACGATG 996
DB 720 ACCTCAAGGACTCCACCTATGACGACATACCGTGGGGGGGACTGACGATG 770

RESULT 6
BG912321 799 bp mRNA EST 05-JUN-2001
LOCUS 602806864F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4939239
DEFINITION 5', mRNA sequence.
ACCESSION BG912321
VERSION BG912321.1 GI:14292797
KEYWORDS EST.

SOURCE
ORGANISM

human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 799)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM10875 row: o column: 16
High quality sequence stop: 768.

location/Qualifiers

source

1..799

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4939239"

/clone_lib="NCI_CGAP_Brn67"

/tissue_type="anaplastic oligodendroglioma with lp/19q
loss"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.3 Kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 163 a 238 c 251 g 147 t

ORIGIN

Query Match 17.3%; Score 684; DB 11; Length 799;
Best Local Similarity 95.7%; Pred. No. 7.8e-117;
Matches 758; Conservative 0; Mismatches 25; Indels 9; Gaps 5;

QY 2443 ACTACATGAGAACGGCGACCTCAACCACTTCTCTAGTGGCCACCACTGAGGACAAG 2502

DB 1 ACTACATGAGAACGGCGACCTCAACCACTTCTCTAGTGGCCACCACTGAGGACAAG 60

QY 2503 CAGCCGAGGGGCCCTTGGGACGGCGAGCTGCGCAGGGGCCACCATCAGCTACCCAA 2562

DB 61 CAGCCGAGGGGCCCTTGGGACGGCGAGCTGCGCAGGGGCCACCATCAGCTACCCAA 120

QY 2563 TGTCTGTCATGTGGCAGCCCGAGATGCGCTCCGGCATGCGCTATCTGGCCACACTCAACT 2622

DB 121 TGTCTGTCATGTGGCAGCCCGAGATGCGCTCCGGCATGCGCTATCTGGCCACACTCAACT 180

QY 2623 TTGTATACATGCGGACCTGCGCAGCGGAACTGCTAGTTGGGGAATAATTTACCATCAAA 2682

DB 181 TTGTATACATGCGGACCTGCGCAGCGGAACTGCTAGTTGGGGAATAATTTACCATCAAA 240

QY 2683 TGGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACATATTACCGTGTGACGGCC 2742

DB 241 TGGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACATATTACCGTGTGACGGCC 300

QY 2743 GGGCAGTGTGCTCCCATCCGCTGGATGGCTGGGAGTGCATCTCATGGGGAAGTTCACGA 2802

DB 301 GGGCAGTGTGCTCCCATCCGCTGGATGGCTGGGAGTGCATCTCATGGGGAAGTTCACGA 360

QY 2803 CTGCCAGTGTGCTGGGCTTTGGTGACCCCTTGGTGTGGGAGGTGCTGCTGTAGGG 2862

DB 361 CTGCCAGTGTGCTGGGCTTTGGTGACCCCTTGGTGTGGGAGGTGCTGCTGTAGGG 420

QY 2863 CCCAGCCCTTTGGGAGCTCACCGAGCAGCGTTCATCCAGAACGCGGGGAGTCTCTTC 2922

DB 421 CCCAGCCCTTTGGGAGCTCACCGAGCAGCGTTCATCCAGAACGCGGGGAGTCTCTTC 480

QY 2923 GGGACAGGGCGGCGAGGTGTACTGTCCCGCCGCCCTGCCTGCCCGC-AGGSCCTATAT 2981


```

Db 481 GGGACACGGCCGGCAGGTACCTGTCTCCGGCGCGCTGCTGCCCAAGGGCCTATAT 540
      |||
QY 2982 GAGCTGATGCTTCGGTGTGGAGCGGAGTCTGAGCAGGACACCCCTTTTCCAGCTG 3041
      |||
Db 541 GAGCTGATGCTTCGGTGTGGAGCGGAGTCTGAGCAGGACCA-CCTTTTCAGCTG 599
      |||
QY 3042 CATCGGTTCTGTCAGAGGATGACCTCAACACCGTGTGTAATCACACATCCAGCTG 3096
      |||
Db 600 CATCGGTTCTGTCAGAGGATGACCTCAACACCGTGTGTAATCACACATCCAGCTG 659
      |||
QY 3097 CCCTCCCTCAGGAGTATCCAGGGAAGCC-AGTGACACTAAA-CARAGAGACACAA 3154
      |||
Db 660 CCCTCCCTCAGGAGGATGATCCAGGGAAGCCAAAGTACACTAAAACCAAGAGACACAA 719
      |||
QY 3155 TGCACTCTGCCCTTCCCTCCCTGAGACGCCCATCACCCTAATAGAGCAGTGAGACTG 3214
      |||
Db 720 TGGGACTCTGCCCTTCCCTCCCGAGGCGCATCAACTCTAATAGAGCGGTGAGATGA 779
      |||
QY 3215 CAGGTGGGCTG 3226
      |||
Db 780 CAGGTGGTGGG 791
      |||

RESULT 7
BE744806
LOCUS 601575025F1 NTH_MGC_9 Homo sapiens cDNA clone IMAGE:3856046 5',
DEFINITION mRNA sequence.
ACCESSION BE744806
VERSION BE744806.1 GI:10158798
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
Plate: LCM519 row: a column: 07
High quality sequence stop: 713.
FEATURES
Location/Qualifiers
1..715
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3836046"
/clone_lib="NTH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: ovary; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 130 a 226 c 212 g 147 t
ORIGIN

```

Query Match 17.2%; Score 680.4; DB 10; Length 715;
 Best Local Similarity 99.4%; Pred. No. 3.6e-116;
 Matches 714; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

```

QY 2682 ATGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACTATTTACCTGTGTCAGGGC 2741
      |||
Db 1 ATGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACTATTTACCTGTGTCAGGGC 59
      |||
QY 2742 CGGCGAGTGTGCTCCCTGATGCTGGAGTGCATCCTCATGCGGGAAGTTTCACG 2801
      |||
Db 60 CGGCGAGTGTGCTCCCTGATGCTGGAGTGCATCCTCATGCGGGAAGTTTCACG 119
      |||
QY 2802 ACTGCGAGTGTGCTGGGCTTTGCTGACCCCTGCGGAGTGTGCTGCTCTGCTGAGG 2861
      |||
Db 120 ACTGCGAGTGTGCTGGGCTTTGCTGACCCCTGCGGAGTGTGCTGCTGCTGAGG 179
      |||
QY 2862 GCCCAGCCCTTTGGGAGCTCACCGAGGAGCAGTATCAGAGAACGGGGGAGTTCTTC 2921
      |||
Db 180 GCCCAGCCCTTTGGGAGCTCACCGAGGAGCAGTATCAGAGAACGGGGGAGTTCTTC 239
      |||
QY 2922 CGGCGACAGGCGCGGAGGTGTACCTGCTCCCGCGCTGCCCTGCCCGCAGGCGCTATAT 2981
      |||
Db 240 CGGCGACAGGCGCGGAGGTGTACCTGCTCCCGCGCTGCCCTGCCCGCAGGCGCTATAT 299
      |||
QY 2982 GAGCTGATGCTTCGGTGTGCGGAGCTGTGAGCAGGACACACCTTTTCCAGCTG 3041
      |||
Db 300 GAGCTGATGCTTCGGTGTGCGGAGCTGTGAGCAGGACACACCTTTTCCAGCTG 359
      |||
QY 3042 CATCGGTTCTGTCAGAGGATGACCTCAACACCGTGTGTAATCACACATCCAGCTG 3101
      |||
Db 360 CATCGGTTCTGTCAGAGGATGACCTCAACACCGTGTGTAATCACACATCCAGCTG 419
      |||
QY 3102 CCCTCAGGAGTGTCCAGGGAAGCCAGTGACACTAAACAAAGAGGACACATGGCACC 3161
      |||
Db 420 CCCTCAGGAGGAGTGTGCA-GGGAAGCCAGTGACACTAAACAAAGAGGACACATGGCACC 478
      |||
QY 3162 TCTGCCCTTCCCTCCCGACAGCCCATCACTCTAATAGAGGAGTGTGAGTGTGCTG 3221
      |||
Db 479 TCTGCCCTTCCCTCCCGACAGCCCATCACTCTAATAGAGGAGTGTGAGTGTGCTG 538
      |||
QY 3222 GCTGGGCCACACCCAGGAGTGTGATGCCCTTCTCCCTTCTGCGACACTCTCATGTCC 3281
      |||
Db 539 GCTGGGCCACACCCAGGAGTGTGATGCCCTTCTCCCTTCTGCGACACTCTCATGTCC 598
      |||
QY 3282 CTTCTCTGCTTCTCTCTAGAGCCCTGTCGCCACCCAGCTGCTGTGGATGG 3341
      |||
Db 599 CTTCTCTGCTTCTCTCTAGAGCCCTGTCGCCACCCAGCTGCTGTGGATGG 657
      |||
QY 3342 ATCTCTCCACCTCTCTAGCCATCCCTTGGGAAGGTTGGGGAATATAGGATA 3399
      |||
Db 658 ATCTCTCCACCTCTCTAGCCATCCCTTGGGAAGGTTGGGGAATATAGGATA 715
      |||

RESULT 8
BF345815
LOCUS 602017891F1 NCI-GAP_Brn67 Homo sapiens cDNA clone IMAGE:4153367
DEFINITION 5', mRNA sequence.
ACCESSION BF345815
VERSION BF345815.1 GI:11293410
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLM9421 row: f column: 24

High quality sequence stop: 719.

Location/Qualifiers

1..719

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4153367"

/clone_lib="NCI_CGAP_Brn67"

/tissue_type="anaplastic oligodendroglioma with lp/19q loss"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

132 a 217 c 230 g 140 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 717; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 2397 CTGCTGGCGCTGTGTGTCAGGACGACCCCTCTGCATGATTACTACTACATGAGAAC 2456

Db 1 CTGCTGGCGCTGTGTGTCAGGACGACCCCTCTGCATGATTACTACTACATGAGAAC 60

QY 2457 GCGACCTCAACAGTTCTCAGTGGCCACACAGCTGGAGGACAAAGGAGCGGGGCC 2516

Db 61 GCGACCTCAACAGTTCTCAGTGGCCACACAGCTGGAGGACAAAGGAGCGGGGCC 120

QY 2517 CTGCGGACGCGAGCTGCGAGGCGCCACCATCATGCTACCAATGCTGCTCATGTG 2576

Db 121 CTGCGGACGCGAGCTGCGAGGCGCCACCATCATGCTACCAATGCTGCTCATGTG 180

QY 2577 GCAGCCACAGATCGCTCGGATCGCTATCTGGCCACACTCAACTTGTACATCGGGAC 2636

Db 181 GCAGCCACAGATCGCTCGGATCGCTATCTGGCCACACTCAACTTGTACATCGGGAC 240

QY 2637 CTGCGGACGCGAGCTGCTAGTTGGGAAATTTACCAATCAAAATCGCAGACTTTGGC 2696

Db 241 CTGCGGACGCGAGCTGCTAGTTGGGAAATTTACCAATCAAAATCGCAGACTTTGGC 300

QY 2697 ATGAGCGGAAACCTTATGCTGGGACTATTACGCTGTGAGGCGCGGAGCTGCTGCC 2756

Db 301 ATGAGCGGAAACCTTATGCTGGGACTATTACGCTGTGAGGCGCGGAGCTGCTGCC 360

QY 2757 ATCCGCTGGATGCGCTGGGAGTGCATCTCATGGGAAAGTTCACGACTGCGAGTGCAGTG 2816

Db 361 ATCCGCTGGATGCGCTGGGAGTGCATCTCATGGGAAAGTTCACGACTGCGAGTGCAGTG 420

QY 2817 TGGGCTTTGGTGTGACCTGTGGAGTGCCTGATGCTGTAGGCGCCAGCCCTTTGGG 2876

Db 421 TGGGCTTTGGTGTGACCTGTGGAGTGCCTGATGCTGTAGGCGCCAGCCCTTTGGG 480

QY 2877 CAGCTCACCGACGAGCTCATCGAGACGCGGGAGTCTTCCGGACACAGGCGCG 2936

Db 481 CAGCTCACCGACGAGCTCATCGAGACGCGGGAGTCTTCCGGACACAGGCGCG 539

QY 2937 CAGGTGTACCTGTCCGGCGCCCTGCTGCCCGCAGGCGCTATATGAGTGTGCTTCGG 2996

Db 540 CAGGTGTACCTGTCCGGCGCCCTGCTGCCCGCAGGCGCTATATGAGTGTGCTTCGG 599

QY 2997 TGTGGAGCGGAGTGTGAGCAGCAGCACCCCTTTTCCAGCTGCATCGGTTCTTGGCA 3056

Db 600 TGTGGAGCGGAGTGTGAGCAGCAGCACCCCTTTTCCAGCTGCATCGGTTCTTGGCA 658

QY 3057 GAGGATGCACTCAACA-CGGTGTGATCAGATCATCATGCTGCCCTCCCTCAGGAGTGA 3115

Db 659 GAGGATGCACTCAACA-CGGTGTGATCAGATCATCATGCTGCCCTCCCTCAGGAGTGA 717

OY 3116 TC 3117

Db 718 TC 719

RESULT 9

BI193181

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Cloning by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLM852 row: b column: 16

High quality sequence stop: 675.

Location/Qualifiers

1..744

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5090223"

/clone_lib="NIH_MGC_42"

/tissue_type="epithelioid carcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. 1"

149 a 212 c 229 g 154 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 714; Conservative 0; Mismatches 15; Indels 3; Gaps 3;

QY 2237 TTTCCCTTAATGTGCGTAAGGACACCCCTTTGCTGCTAGCTGCTCAAGATCTTACGCCC 2296

Db 2 TTTCCCTTAATGTGCGTAAGGACACCCCTTTGCTGCTAGCTGCTCAAGATCTTACGCCC 61

QY 2297 AGATGCCACCAAGATGCCAGCTTCTCTCCAGGAATGATTCCTGAAAGAGGT 2356

Db 62 AGATGCCACCAAGATGCCAGCTTCTCTCCAGGAATGATTCCTGAAAGAGGT 121

QY 2357 GAAGATCATGTCCAGCTCAAGGACCCCAACATCATTCGGCTGCTGGCGGTGTGTGCA 2416

Db 122 GAAGATCATGTCCAGCTCAAGGACCCCAACATCATTCGGCTGCTGGCGGTGTGTGCA 181

QY 2417 GGACGACCCCTCTGCTGATGATTACTGACTACATGGAGAGCGGAGCTCAACAGTCTCT 2476

Db 182 GGACGACCCCTCTGCTGATGATTACTGACTACATGGAGAGCGGAGCTCAACAGTCTCT 241

QY 2477 CAGTCCACAGCTGAGGACAAAGGAGCGAGGGGCGCCCTGGGACGGCAGGCTGC 2536
 Db 242 CAGTCCACAGCTGAGGACAAAGGAGCGAGGGGCGCCCTGGGACGGCAGGCTGC 301
 QY 2537 GCAGGGCCACCATCAGCTACCAATGCTGCTGATGTGGCAGCCAGATCGCCTCGG 2596
 Db 302 GCAGGGCCACCATCAGCTACCAATGCTGCTGATGTGGCAGCCAGATCGCCTCGG 361
 QY 2597 CATCGCTATCTGGCCACACTCAACTTTGTACATCGGGACCTGGCCACGGGAACCTGCCT 2656
 Db 362 CATCGCTATCTGGCCACACTCAACTTTGTACATCGGGACCTGGCCACGGGAACCTGCCT 421
 QY 2657 AGTTGGGAAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATGC 2716
 Db 422 AGTTGGGAAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATGC 481
 QY 2717 TGGGGAATATTACCGTGTGAGGCGCGGAGTGTGCTGCCCATCGCTGGATGGCTGGGA 2776
 Db 482 TGGGGAATATTACCGTGTGAGGCGCGGAGTGTGCTGCCCATCGCTGGATGG-CTGGGA 540
 QY 2777 GTGATCCTCATGGGAAATTCAGACTGCGAGTGTGAGTGTGGGCTTTGGTGTGACCCCT 2836
 Db 541 GTGATCCTCATGGGAAATTCAGACTGCGAGTGTGAGTGTGGGCTTTGGTGTGACCCCT 600
 QY 2837 GTGGAGGTGCTGATGCTGTGTAGGCGCCAGCCCTTTGGGACGCTCACCGACGAGCAGGT 2896
 Db 601 GTGGAGGTGCTGATGCTGTGTAGGCGCCAG -CCTTTGGGCAAGTCACCGACGAGCAGGT 659
 QY 2897 CATCGGAAAG -CGGGGGAGTCTTCGGGACGAGCGCGGAGGTGTACCTGTCCCGGC 2955
 Db 660 CATCGGAAAGCGGGGGAGTCTTCGGGACGAGCGCGGAGGTGTACCTGTCCCGGC 719
 QY 2956 CGCCTGCCTGCC 2967
 Db 720 CGCCTGCCTGCC 731

RESULT 10
 LOCUS BG696171 718 bp mRNA EST 07-MAY-2001
 DEFINITION 602659285F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4802421 5', mRNA sequence.
 ACCESSION BG696171
 VERSION BG696171.1 GI:13961044
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL) DNA
 Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10696 row: j column: 22
 High quality sequence stop: 717.
 Location/Qualifiers
 1. 718
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4802421"
 /clone_lib="NCI_CGAP_Skn3"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library.
 BASE COUNT 133 a 228 c 210 g 147 t
 ORIGIN

Query Match 16.7%; Score 662.6; DB 11; Length 718;
 Best Local Similarity 98.7%; Pred. No. 7.2e-113;
 Matches 710; Conservative 0; Mismatches 4; Indels 5; Gaps 4;
 QY 2670 TTCACCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACTATTAC 2729
 Db 1 TTCACCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACTATTAC 60
 QY 2730 CGTGTGAGGCGCGGCGAGTGTGCCATCCGCTGGATGGCCTGGGAGTGCCTCATG 2789
 Db 61 CGTGTGCA -GGCCGGGCGAGTGTGCCATCCGCTGGATGGCCTGGGAGTGCCTCATG 119
 QY 2790 GGGAAGTTCACGACTGCGAGTGTGGGCTTTGGTGTGACCCCTGTGGGAGTGTCTG 2849
 Db 120 GGGAAGTTCACGACTGCGAGTGTGGGCTTTGGTGTGACCCCTGTGGGAGTGTCTG 179
 QY 2850 ATGCTCTGTAGGCGCCAGCCCTTTGGCAGCTCACCCAGCAGCAGTGCATCGAAGCGC 2909
 Db 180 ATGCTCTGTAGGCGCCAGCCCTTTGGCAGCTCACCCAGCAGCAGTGCATCGAAGCGC 239
 QY 2910 GGGAGTTCCTCCGGGACGAGCGCGGAGTGTACCTGTCCCGGCGCCCTGCTGCCCG 2969
 Db 240 GGGAGTTCCTCCGGGACGAGCGCGGAGTGTACCTGTCCCGGCGCCCTGCTGCCCG 299
 QY 2970 CAGGCGCTATATGAGTGTGCTCGTGTGGGCGGAGTGTGAGCAGCGGACCC 3029
 Db 300 CAGGCGCTATATGAGTGTGCTCGTGTGGGCGGAGTGTGAGCAGCGGACCC 359
 QY 3030 TTTTCCAGCTGCATCGGTTCTTGGCAGAGGATGCATCAACAGGTTGTAATACACAT 3089
 Db 360 TTTTCCAGCTGCATCGGTTCTTGGCAGAGGATGCATCAACAGGTTGTAATACACAT 419
 QY 3090 CCAGCTGCCCTCCCTCAGGAGTGTATCCAGGGAGCCAGTGCATCAAAACAAGAGGA 3149
 Db 420 CCAGCTGCCCTCCCTCAGGAGTGTATCCAGGGAGCCAGTGCATCAAAACAAGAGGA 479
 QY 3150 CACAATGGCACCTCTGCCCTTCCCTCCCGACAGCCCATCACCTCTAATAGAGCAGTGA 3209
 Db 480 CACAATGGCACCTCTGCCCTTCCCTCCCGACAGCCCATCACCTCTAATAGAGCAGTGA 539
 QY 3210 GACTGCAGTGTGGCTGGGCGCCACCCAGGAGTGTATGCCCTTCT -CCCTTCTCTGGAC 3267
 Db 540 GACTGCAGTGTGGCTGGGCGCCACCCAGGAGTGTATGCCCTTCTACCCATTCCTGGAC 599
 QY 3268 ACACCTCTATGTCCTT -CCTGTTCTTCTCTAGAGCCCTGTGCGCCACCCAGCT 3326
 Db 600 ACACCTCTATGTCCTTACTTACTTCTCTAGAGCCCTGTGCGACCCACCCAGCT 659
 QY 3327 GGTCTGTGGATGGGATCTCTCCACCCCTCTAGCC -ATCCTTGGGGAAGGTGGG 3384
 Db 660 GGTCTGTGGATGGGATCTCTCTCCACCCCTCTAGCCATCTCTGGGGAAGGTGGG 718

RESULT 11

LOCUS AL043251 772 bp mRNA EST 29-FEB-2000
 DEFINITION DKFZp434K0223_r1 434 (synonym: htes3) Homo sapiens cDNA clone
 ACCESSION AL043251
 VERSION AL043251
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 772)
 Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

TITLE EST (Blum, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Blum H
MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DFGZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.

No s1 sequence available.

This clone (DKF2p43K0223) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

Location/Qualifiers

1..772

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="DKF2p43K0223"

/clone_lib="434 (synonym: htes3)"

/tissue_type="testis"

/dev_stage="adult"

/lab_host="DH10B"

/note="vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 118 a 250 c 260 g 143 t

ORIGIN

Query Match 16.7%; Score 660.2; DB 10; Length 772;

Best Local Similarity 92.4%; Pred. No. 2e-112;

Matches 750; Conservative 0; Mismatches 4; Indels 58; Gaps 3;

QY 64 CTCCTGGGAGCGCTCCCGACACCCGAGCCCGCGCGCTCCCGCTCCCGCTCCCG 123

DB 1 CTCCTGGGAGCGCTCCCGACACCCGAGCCCGCGCGCTCCCGCTCCCGCTCCCG 60

QY 124 GCTCTGGCTCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 183

DB 61 GCTCTGGCTCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 120

QY 184 GGTCTGGAGCGCTGGTCTCCGCGGAGAGCGATGAGAGGTGTCTGAAGTGGTATCA 243

DB 121 GGTCTGGAGCGCTGGTCTCCGCGGAGAGCGATGAGAGGTGTCTGAAGTGGTATCA 157

QY 244 CTGAGCGATGGGTTGGACTTGAAGGAATGCCAAGAGATGTCGCCACCCCTTAGGC 303

DB 158 -----AGATGCTGCCGCCACCCCTTAGGC 184

QY 304 CCGAGGATCAGGAGCTATGGACACAGGCGCTGTCTACTTCTGTCTGTCTGTGG 363

DB 185 CCGAGGATCAGGAGCTATGGACACAGGCGCTGTCTACTTCTGTCTGTCTGTGG 244

QY 364 TGGCAAGTGGAGTGTGACATGAAGGAGACATTTTGTATCTGCAAGTCCGCTATGCC 423

DB 245 TGGCAAGTGGAGTGTGACATGAAGGAGACATTTTGTATCTGCAAGTCCGCTATGCC 304

QY 424 TGGCATGCAGGACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCTGTGTCAG 483

DB 305 TGGCATGCAGGACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCTGTGTCAG 364

QY 484 ATTCCATCGCGCGCCACAGCAGGTTGAGAGCAGTGCAGGGGATGGGGCTGTGGCC 543

DB 365 ATTCCATCGCGCGCCACAGCAGGTTGAGAGCAGTGCAGGGGATGGGGCTGTGGCC 424

QY 544 CCGCAGGTCGCTGTTCCTCCAGGAGGAGTACTTGCAGTGGATCTACACAGTCTCC 603

DB 425 CCGCAGGTCGCTGTTCCTCCAGGAGGAGTACTTGCAGTGGATCTACACAGTCTCC 484

QY 604 ACCTGTGGCTCTGTGGGACCCAGGACCGCATGCCGGGGCTTGGGCAAGGATCTT 663

DB 485 ACCTGTGGCTCTGTGGGACCCAGGACCGCATGCCGGGGCTTGGGCAAGGATCTT 544

QY 664 CCGGAGCTACCGCTCGGTTACTCCCGGATGGTCGCGCTGGTGGCTGAAGGACC 723
DB 545 CCGGAGCTACCGCTCGGTTACTCCCGGATGGTCGCGCTGGTGGCTGAAGGACC 604
QY 724 GCTGGGTTCAGGAGTGTATCTCAGCAATGAGGACCCCTGAGGAGTGGTCTGAAGGACC 783
DB 605 GCTGGGTTCAGGAGTGTATCTCAGCAATGAGGACCCCTGAGGAGTGGTCTGAAGGACC 664
QY 784 TTGGCCCCCATGTTCCCGACTGGTTCGCTTACCCCGGCTGACCCGGGTCTACGA 843
DB 665 TTGG-CCCCCATGTTCCCGACTGGTTCGCTTACCCCGGCTGACCCGGGTCTACGA 723
QY 844 GTGCTGTCTCGGAGTAGAGTCTATGCTGC 875
DB 724 GCGTCTGTCTCGGAGTAGA-CTCTATGCTGC 754

RESULT 12

BE304984/c

LOCUS

DEFINITION

BE304984 912 bp mRNA EST. 13-JUL-2000

600942930T2 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:2959615 3',

mRNA sequence.

ACCESSION BE304984

VERSION BE304984.1 GI:9176727

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 912)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: rgs@nih.gov

Tissue procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: NIH Intramural Sequencing Center

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LCM51 row: k column: 08

High quality sequence start: 9

High quality sequence stop: 702.

FEATURES

Location/Qualifiers

1..912

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2959615"

/clone_lib="NIH_MGC_15"

/tissue_type="adenocarcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-CDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 211 a 244 c 277 g 180 t

ORIGIN

Query Match 16.5%; Score 655.6; DB 10; Length 912;

Best Local Similarity 95.4%; Pred. No. 1.4e-111;

Matches 740; Conservative 0; Mismatches 29; Indels 7; Gaps 6;

QY 3098 CCTCTCCTCAGGAGTGTATCCAGGGAGCCAGTGCAC--TAAACAAGAGGACAAAT 3155

DB 778 CCTCTCCTCAGGAGTGTATCCAGGGAGCCAGTGCAC--TAAACAAGAGGACAAAT 719.

QY 3156 GGCACCTCTGCCCTTCCCTCCCGACAGCCATCACCTCTTAATAGAGGAGTGCAGTGC 3215

Db 718 CGGCACTCGCGCGCTCCCTCCGACAGCCCATCACTCTAATAGAGTCAGGGAGACCG 659
QY 3216 AGTGGGCTGGGCGCCACCCAGGAGCTGATGCCCTTCTCCCTTCTGGACACATCTC 3275
Db 658 AGTGGGCTGGGCGCCACCCAGGAGTGA-GGCCCTTCCCTCCCTGGACACATCTC 600
QY 3276 ATGTGCCC-TTCTCTGTTC-TCTTCTAGAAAGCCCTGTGCGCCACCCAGCTGGCTCG 3333
Db 599 AGTCCCCCTCTCTGTCTCTCTAGAGCCCTGTGCGCCACCCAGCTGGCTCG 540
QY 3334 TGGATGGGATCTCTCCACCCCTCTAGCCATCCCTTGGGAGAGGTGGGAGAAATAT 3393
Db 539 TGGATGGGATCTCTCCACCCCTCTAGCCATCCCTTGGGAGAGGTGGGAGAAATAT 480
QY 3394 AGGATAGACACTGGACATGGC-QATTGAGACACTGGGCGCCACCTGGACACACTGATT 3452
Db 479 AGGATAGACACTGGACATGGCCTATTGGAGACCTGGGCGCCACCTGGACACACTGATT 420
QY 3453 CTTGAGAGGTGGCTGCG-CCGACGCTCTCTCTCCCTGTGACACACTGGACCCACTGG 3511
Db 419 CTTGAGAGGTGGCTGCGGCCCCAGCTTCTCTCTCTCTGACACACTGGACCCACTGG 360
QY 3512 CTGAGAATCTGGGGTGGAGGAGCAAGAGAGGAGAAATGTTTCTCTGTGCTGCTC 3571
Db 359 CTGAGAATCTGGGGTGGAGGAGCAAGAGAGGAGAAATGTTTCTCTGTGCTGCTC 300
QY 3572 CTGTACTTGTCTAGCTTGGGCTTCTCTCTCTCCATCACTGAAACACTGGACCTGG 3631
Db 299 CTGTACTTGTCTAGCTTGGGCTTCTCTCTCTCCATCACTGAAACACTGGACCTGG 240
QY 3632 GGTAGCCCGCCGAGCCCTAGTCACCCCTCCCTCCCTCCATCACTGAAACACTGGAGTAA 3691
Db 239 GGTAGCCCGCCGAGCCCTAGTCACCCCTCCCTCCCTCCATCACTGAAACACTGGAGTAA 180
QY 3692 CTTCTCTAAGCCTATACGTTCTCTGTGGAGTAAATATTGGGATGGGGGAAAGAGGAGC 3751
Db 179 CTTCTCTAAGCCTATACGTTCTCTGTGGAGTAAATATTGGGATGGGGGAAAGAGGAGC 120
QY 3752 AACGCCCATAGCCTTGGGTTGGACATCTCTAGTGTAGCTGCCCATGATTTTCTAT 3811
Db 119 AACGCCCATAGCCTTGGGTTGGACATCTCTAGTGTAGCTGCCCATGATTTTCTAT 60
QY 3812 ATACTTGGGTTTGTACATTTTGGGGGAGACACAGATTTTACATAA 3867
Db 59 AATCACTTGGGTTTGTACATTTTGGGGGAGACACAGATTTTACATAA 4

RESULT 13
BF944646
LOCUS BF944646 662 bp mRNA EST 22-JAN-2001
DEFINITION PMO-NN1171-181000-001-c05 NN1171 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF944646
VERSION BF944646.1 GI:12361921
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 662)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PMO&t2=PMO-NN1171-
181000-001-c05&t3=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 660.

FEATURES
source

1..662
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NN1171"
/dev_stage="Adult"
/note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (O.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 124 a 214 c 191 g 132 t 1 others
ORIGIN

Query Match 16.5%; Score 654.6; DB 11; Length 662;
Best Local Similarity 99.2%; Pred. No. 2.2e-111; Indels 0; Gaps 0;
Matches 657; Conservative 0; Mismatches 5;
QY 2798 CACGACTGGAGTGACGTGGGCTTTGGTGACCTGTGGAGGTGCTGATGCTCTG 2857
Db 1 CACGACTGGAGTGACGTGGGCTTTGGTGACCTGTGGAGGTGCTGATGCTCTG 60
QY 2858 TAGGGCCAGCCCTTTGGGAGCTCACGACGAGAGGTCTCGAAGACGGGGGAGTT 2917
Db 61 TAGGGCCAGCCCTTTGGGAGCTCACGACGAGAGGTCTCGAAGACGGGGGAGTT 120
QY 2918 CTTCCGGGACAGGGCGGCGAGGTGTACCTGTCCGCGCGCTGCTGCCGAGGCGCT 2977
Db 121 CTTCCGGGACAGGGCGGCGAGGTGTACCTGTCCGCGCGCTGCTGCCGAGGCGCT 180
QY 2978 ATATGAGCTGATCTTCCGTGCTGGAGCGGGAGTGTGAGCAGCAGCCCTTTTCCCA 3037
Db 181 ATATGAGCTGATCTTCCGTGCTGGAGCGGGAGTGTGAGCAGCAGCCCTTTTCCCA 240
QY 3038 GCTGCATCTGGTCTCGGAGAGATGCATCAACACGGTGTGAATCACACATCCAGCTGC 3097
Db 241 GCTGCATCTGGTCTCGGAGAGATGCATCAACACGGTGTGAATCACACATCCAGCTGC 300
QY 3098 CCTCCTCCTCAGGAGTGATCCAGGGGAAGCCAGTGACACTAAACAGAGGACACAATGG 3157
Db 301 CCTCCTCCTCAGGAGTGATCCAGGGGAAGCCAGTGACACTAAACAGAGGACACAATGG 360
QY 3158 CACCTGTGCTTCCCTCCGACAGCCCATCACTCTAATAGAGCAGTGAGACTGCAG 3217
Db 361 CACCTGTGCTTCCCTCCGACAGCCCATCACTCTAATAGAGCAGTGAGACTGCAG 420
QY 3218 GTGGGCTGGGCGCCACCCAGGAGCTGATGCCCTTCTCCCTTCTCGGACACACTCTCAT 3277
Db 421 GTGGGCTGGGCGCCACCCAGGAGCTGATGCCCTTCTCCCTTCTCGGACACACTCTCAT 480
QY 3278 GTCCCTTCTCTCTTCTCTCTAGAAAGCCCTGTGCGCCACCCAGCTGGCTGTGA 3337
Db 481 GTCCCTTCTCTCTTCTCTCTAGAAAGCCCTGTGCGCCACCCAGCTGGCTGTGA 540
QY 3338 TGGGATCTCTCCACCCCTCTCTAGCCATCCCTTGGGGAGGCTGGGAGAAATATAGA 3397
Db 541 TGGGATCTCTCCACCCCTCTCTAGCCATCCCTTGGGGAGGCTGGGAGAAATATAGA 600
QY 3398 TAGACACTGGACATGGCCCATTTGGAGCAGCTGGGCCCCACTGGACAACTGATTCCTGG 3457

```

|||||
Db 601 TAGACACTGGACATGGGCGCATTTGAGCACCTTGGGCCCGCCACTGGACACACTGATTCCTGG 660
Qy 3458 AG 3459
Db 661 AG 662

RESULT 14
BE249888 877 bp mRNA EST 13-JUL-2000
LOCUS 600942930F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:2959615 5',
DEFINITION mRNA sequence.
ACCESSION BE249888
VERSION BE249888.1 GI:91119991
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 877)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM51 row: k column: 08
High quality sequence stop: 725.
FEATURES
Location/Qualifiers
1..877
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2959615"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 167 a 248 c 282 g 180 t
ORIGIN

Query Match 16.5%; Score 653.8; DB 10; Length 877;
Best Local Similarity 91.0%; Pred. No. 3e-111;
Matches 782; Conservative 0; Mismatches 47; Indels 30; Gaps 7;

Qy 2082 GTGCTGCACATGCCCGCCAGGGGCGATCGGGATGGGCCCGCCAGAGTGGATTTCCTCGA 2141
Db 1 GTGCTGCACATGCCCGCCAGGGGCGATCGGGATGGGCCCGCCAGAGTGGATTTCCTCGA 60

Qy 2142 TCTGCATCCGCTCAAGGAGAAGTTGGCGAGGGCCAGTTGGGGAGGTGCACCTGTGT 2201
Db 61 TCTGCATCCGCTCAAGGAGAAGTTGGCGAGGGCCAGTTGGGGAGGTGCACCTGTGT 120

Qy 2202 GAGGTGCGACAGCCCTCAAGATCTGTGATGTC-TTGATTCCCGCTTAATGTGCGTAGGG 2260
Db 121 GAGGTGCGACAGCCCTCAAGATCTGTGATGTC-TTGATTCCCGCTTAATGTGCGTAGGG 180

Qy 2261 ACACCCCTTGTGTTAGTGTGTCAAGATCTTACGGCCAGATGCCACCAAGAGTGCACGTT 2320
Db 181 ACACCCCTTGTGTTAGTGTGTCAAGATCTTACGGCCAGATGCCACCAAGAGTGCACGTT 233

```

```

Qy 2321 CTCCTTGTCTCCAGGAATGATTTCTCTGAAGAGGTGAAGATCATGTCTGAGGCTCAAGGA 2380
Db 234 -----CCAGGAATGATTTCTCTGAAGAGGTGAAGATCATGTCTGAGGCTCAAGGA 282
Qy 2381 CCCCAACATCATTCGGCTGCTGGCGTGTGTGTGAGGACGACCCCTCTGATGATTAAC 2440
Db 283 CCCAAACATCATTCGGCTGCTGGCGTGTGTGTGAGGACGACCCCTCTGATGATTAAC 342
Qy 2441 TGACTACATGAGAACGGCGACCTCAACAGTTCTCTAGTCCCGCCAGCTGGAGGACAA 2500
Db 343 TGACTACATGAGAACGGCGACCTCAACAGTTCTCTAGTCCCGCCAGCTGGAGGACAA 402
Qy 2501 GGCAGCGAGGGGGCCCTTGGGAGCGGCGAGGCTGCGCAGGGGCCACCATCAGTACCC 2560
Db 403 GGCAGCGAGGGGGCCCTTGGGAGCGGCGAGGCTGCGCAGGGGCCACCATCAGTACCC 462
Qy 2561 AATGCTCTGCATGTGGCAGCCCGCAGATGCGCTCCGGCATGCGCTATCTGGCCACACTCAA 2620
Db 463 AATGCTCTGCATGTGGCAGCCCGCAGATGCGCTCCGGCATGCGCTATCTGGCCACACTCAA 522
Qy 2621 CTTTGTACATCGGGACCTGGCCACGCGAACTGCTAGTTGGGAAATTTCCACATCAA 2680
Db 523 CTTTGTACATCGGGACCTGGCCACGCGAACTGCTAGTTGGGAAATTTCCACATCAA 582
Qy 2681 AATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACATATTACCGTGTGCA--G 2738
Db 583 AATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACATATTACCGTGTGCACTG 642
Qy 2739 GCGCGGCACTGTCGCCATCGCTGATGCGCTGGAGTGCATCTCATGGG-AACTT 2797
Db 643 GTCGGCGAGTGTGTCCTATCGCTGGATGGCTGGGAATGCTCCCTCATGGGAAAGTT 702
Qy 2798 CAGCACTGCGAGTACGTGTGGCCCTTTGGTGTGACCCCTTGGGAGGTGCTGATGCTCTG 2857
Db 703 CAGCACTGCGAGTACGTGTGGCCCTTT--GGTGTGCGCTGTGGGAGGTGCTGATGCTCT 760
Qy 2858 TAGGCGCCAGCCCTTTGGGCGAGCTCACCGACGACGAGTGCATCGAGAACGCGGGGAGTT 2917
Db 761 GTAGGCCAGCC--TTTGGCAGTCCACGACGAGCGG---TCCCGAGACCCGGGGAGT 814
Qy 2918 CTTCCGGGACCGGCGCG 2936
Db 815 CTTCCGGGACCGGCGCAGG 833

RESULT 15
BE737252 957 bp mRNA EST 15-SEP-2000
LOCUS 601305802F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3640248 5',
DEFINITION mRNA sequence.
ACCESSION BE737252
VERSION BE737252.1 GI:10151244
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 957)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM344 row: c column: 01
High quality sequence start: 22

```

High quality sequence stop: 699.
Location/Qualifiers
1. .957

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3640248"
/clone_lib="NIH_MCC_39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pORF7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."

BASE COUNT 137 a 301 c 318 g 201 t
ORIGIN

Query Match 16.5%; Score 652; DB 10; Length 957;

Best Local Similarity 93.5%; Pred. No. 6.5e-111;

Matches 757; Conservative 0; Mismatches 45; Indels 8; Gaps 7;

```
QY 1023 GTGGTGGGCTGGATGACTTTAGGAAGAGTCAGGAGCTGCGGCTGCGCCAGGCTATGAC 1082
DB 8 GTGGGCTGTGTATGACTTTAGGAAGAGTCAGGAGCTGCGGCTGCGCCAGGCTATGAC 67

QY 1083 TATGTGGATGGAGCAACCAAGCTTCTCCAGTGCTATGAGATGGAGTTTGAGTTT 1142
DB 68 TATGTGGATGGAGCAACCAAGCTTCTCCAGTGCTATGAGATGGAGTTTGAGTTT 127

QY 1143 GACCGGCTGAGGGCTTCCAGGCTATGAGGCTCCACTGTACACATGCACACGCTGGGA 1202
DB 128 GACCGGCTGAGGGCTTCCAGGCTATGAGGCTCCACTGTACACATGCACACGCTGGGA 187

QY 1203 GCCCGTCTGCTGGCGGGTGGAAATGTCGCTTCGGGCTGGCCCTGCCATGGCTGGGAG 1262
DB 188 GCCCGTCTGCTGGCGGGTGGAAATGTCGCTTCGGGCTGGCCCTGCCATGGCTGGGAG 247

QY 1263 GGGAGGCCCATGCGCCCAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGGCTGTC 1322
DB 248 GGGAGGCCCATGCGCCCAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGGCTGTC 307

QY 1323 TCAGTGCCCTTGGCGGCGTGGCTGCGCTTTCAGAGTCCGCTTCCTTTGCGGGG 1382
DB 308 TCAGTGCCCTTGGCGGCGTGGCTGCGCTTTCAGAGTCCGCTTCCTTTGCGGGG 367

QY 1383 CCCTGGTTACTTTCAGGAAATCTCCTTCATCTCTGATGTGGTGAACAATTCCTCTCG 1442
DB 368 CCCTGGTTACTTTCAGGAAATCTCCTTCATCTCTGATGTGGTGAACAATTCCTCTCG 427

QY 1443 GCATGGGAGGACACTTCCCGCCAGCCCTGTGTGGCGGCTGGCCACCTCCACCAAC 1502
DB 428 GCATGGGAGGACACTTCCCGCCAGCCCTGTGTGGCGGCTGGCCACCTCCACCAAC 487

QY 1503 TTCAGAGCTTGGAGCTGGAGCCAGAGGCCAGAGCCCGGTGGCCAGCCGATGGGA 1560
DB 488 TTCAGAGCTTGGAGCTGGAGCCAGAGGCCAGAGCCCGGTGGCCAGCCGATGGGA 547

QY 1561 GCCCGACCGCATCTCATCGGTGCGCTG -GTGGCCATCATCTGCTGCTGCTCATC 1619
DB 548 GCCCGACCGCATCTCATCGGTGCGCTG -GTGGCCATCATCTGCTGCTGCTCATC 607

QY 1620 ATTGCCCTCATCTCTGGCGGCTGCACTGGCGAGGCTCCTCAGCAAGGCTGACGGAGG 1679
DB 608 ATTGCCCTCATCTCTGGCGG -TGGACTGCCCGAGGCTCCTCAGCAGGCTGAACGGAGG 666

QY 1680 GTGTTGGAAGAGAGCTGAGGTTACCTCTCTG -TCCCTGGGGACACTATCTCATCAA 1738
DB 667 GTGTTGGAAGAGAGCTGAGGTTACCTCTCTGTTCCCTTGGGGACACTATCTCATCA 726
```

```
QY 1739 CAACCGCCAGGTCTCTAGAGAGCCACCCCGTACCAGAGCCCGCGCTCGTGGAAATCC 1798
DB 727 AGACCGGCCAGGTCTTAAAGGAGCCACCCCGTAC--AGGGGCCCGCGCTCGTGGATACC 784

QY 1799 GCGCCACTCGGCTCCCTGTGTGTCCTCCCAATGG 1828
DB 785 GGCCCA-TCCGGTCCCTGTGTGCGCATTTG 813
```

Search completed: December 8, 2001, 20:10:20
Job time: 4419 sec

